

A COMPARISON OF MEASURES OF CORE INFLATION

- Designed to remove transitory changes from aggregate price data, measures of underlying, or “core,” inflation are important tools in the monetary policymaking process.
- Somewhat surprisingly, little consensus has been reached on a preferred measure of U.S. core inflation.
- An evaluation of several proposed measures of U.S. core inflation, including the popular ex food and energy series, finds that no measure consistently dominates the others.
- There is arguably too much variability in the nature and sources of transitory price movements to be captured effectively through the design of any individual measure.
- The general practice of focusing on a measure of core inflation that excludes food and energy does not seem to be justified by the analysis.

Robert Rich is a research officer and Charles Steindel a senior vice president at the Federal Reserve Bank of New York.
<robert.rich@ny.frb.org>
<charles.steindel@ny.frb.org>

1. INTRODUCTION

Central banks differ in their specific inflation objectives and conduct of policy. However, they typically confront the common problem of identifying which price changes are permanent and which are transitory. Because of the lagged effects of monetary policy, mistaking the nature of price changes can be extremely costly. For example, the failure to detect the onset of inflationary pressures may lead to a sustained rise in inflation and ultimately require a more prolonged period of policy tightening. Then again, an overreaction to a temporary increase in inflation may result in an unwarranted slowing, and possible decline, in economic activity. Thus, the ability of central banks to differentiate between permanent and transitory price movements is critical for determining the appropriate prescription for monetary policy.

The importance of gauging the persistence of price changes in a timely manner has led to the development of schemes to filter incoming data on aggregate prices. The filtering schemes are designed to remove transitory price movements and thereby produce a measure of underlying, or “core,” inflation. The most common measure of core inflation is aggregate household inflation excluding the contribution of price

The authors have benefited from the suggestions of attendees at the Bank for International Settlements conference on the Evolving Inflation Process and the Federal Reserve Bank of Cleveland–Federal Reserve Bank of Dallas conference on Price Measurement for Monetary Policy. They thank Adam Ashcraft, Steve Cecchetti, Tim Cogley, and two anonymous referees for valuable comments, and are grateful to Steven Reed of the Bureau of Labor Statistics for supplying detailed methodologically consistent CPI data. They also thank David Bedoll, Michael Strain, and Evan LeFlore for research assistance. The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

changes from food and energy. However, alternative core inflation measures have been proposed. Some of these candidate series include ex energy measures (Clark 2001), weighted median measures (Bryan and Cecchetti 1994), and exponentially smoothed measures (Cogley 2002), with proponents citing the superior properties and performance of the respective series across various dimensions.

The lack of consensus on a preferred measure of core inflation might seem surprising given the importance of this information to policymakers. However, a closer examination

The lack of consensus on a preferred measure of core inflation might seem surprising given the importance of this information to policymakers. However, a closer examination of the evidence reveals little uniformity across the dimensions used to compare the proposed measures.

of the evidence reveals little uniformity across the dimensions used to compare the proposed measures. These dimensions include statistical metrics, such as within-sample regression fit and out-of-sample forecast performance as well as more basic considerations relating to the selection of the sample period for the analysis; the data frequency of the price changes (that is, monthly versus quarterly observations); and even the choice of the price measure used to calibrate the core inflation series. Consequently, it is unclear whether the conflicting evidence in support of various measures of core inflation reflects inherent differences in performance capabilities or a lack of standardization in the evaluation process.

This article provides a systematic evaluation of several proposed measures of U.S. core inflation: the popular ex food and energy series, an ex energy series, a weighted median series, and an exponentially smoothed series. To inform the current debate on this issue, we adopt a general framework for the analysis. Regarding the choice of aggregate price indexes, we focus on inflation measures that are likely goals for U.S. monetary policy, namely the consumer price index (CPI) and the personal consumption expenditure (PCE) index. To evaluate the core measures of inflation, we select criteria that have been used in previous studies: ease of design, a similar mean to the goal inflation series, and an ability to track the trend in the goal inflation series.

We also include the explanatory and forecasting capabilities of the core measures of inflation as criteria, but recognize that

the need to specify an econometric model introduces some discretion into the analysis. In an attempt to mitigate this concern, we adopt a benchmark model that relates future changes in inflation to the transitory component of price changes identified by the candidate series. This specification has the advantage of being not only simple to interpret in this context, but also flexible enough to allow us to incorporate alternative horizons into the analysis. As a further check for robustness, we also examine results over different subsamples.

Taken together, we find that no core measure of inflation consistently dominates the others. Further, the performance of the candidate series differs markedly across the aggregate inflation measures, criteria, and sample periods. This conclusion is unaffected by the addition of simple measures of economic slack to the benchmark model. Therefore, we contend that the specifics of the criteria and methods used likely do not account for the differences in performance capabilities of the candidate series. Rather, we suggest that this conflicting evidence reflects the lack of a consistent pattern in transitory price movements. Namely, there is too much variability in the nature and sources of transitory price movements to be captured effectively through the design of the individual core inflation measures. We argue that this interpretation is consistent with the diversity of previous findings using U.S. data and with the work of Hogan, Johnson, and Lafèche (2001) in the Canadian context and Mankikar and Paisley (2002) in the U.K. context. Both studies similarly conclude that there is no single core inflation measure that performs well across-the-board.

Our inability to identify a clear “best” or “worst” measure of either core CPI or core PCE inflation also has implications for some aspects of policy formulation and discussion. While it would be desirable to rely on a single measure of core inflation to perform a multitude of tasks, the evidence does not offer support for this scenario. Consequently, we cannot identify a compelling analytical reason, on either an ex ante or ex post basis, to concentrate attention on a measure of inflation that excludes food and energy prices.

2. MOTIVATION AND CONCEPTS

Almost all central banks are concerned with, and have some mandate to achieve, price stability. Even when ongoing changes in the aggregate price level are anticipated, however, the changes impose costs on economies. These costs need not be directly related to movements in any type of household price measure; they could stem from systematic changes in the prices of all goods and services produced or purchased, including items bought by businesses and governments.

As a practical matter, inflation goals are often linked to movements in a price measure for goods and services purchased by consumers. One reason for this linkage is that the prices for many capital goods purchased by businesses are extremely difficult to measure,¹ as are those for many products provided by governments (such as public education). Moreover, a broad measure of consumer prices should be reasonably successful in capturing the component of aggregate price movements that may affect economic efficiency.

Another reason why a central bank would be concerned with movements in a household or consumer price measure is that many formal escalation arrangements, notably for wages as well as taxes and government benefits, are connected to indexes of consumer prices.² These arrangements could lead to household price movements affecting the distribution of income as well as government revenues and expenditures. In turn, these shifts could influence employment, investment, and basic fiscal policy decisions, and thereby affect the macroeconomic environment faced by monetary policymakers. Thus, there are pragmatic reasons for central banks to concentrate their attention on the consumer component of inflation.

Given a concern with longer term movements in household price inflation, central banks and private agents need some means by which to gauge current performance vis-à-vis a price inflation objective.³ The main reason to focus on the behavior of a core price measure is the belief that there is a significant amount of transitory noise in the movement of aggregate consumer prices. Filtering out the transitory noise gives a better sense of the underlying trend in prices, and thus a better sense of how a measure of current price changes compares with an explicit or inferred longer term goal.⁴ Accordingly, the role of a core price measure lends itself to being interpreted as a means to an end, with low and stable growth of a core price measure serving as an “intermediate target” of policy rather than as a direct “goal.”⁵ This interpretation might also make clear that a central bank’s decision to downplay certain price changes in the conduct and communication of monetary policy does not

¹One could argue that the ideal aggregate inflation index would not include the acquisition prices of capital goods, but rather would include the current “user cost” of existing capital. Nonetheless, as is the case for capital goods acquisition prices, it is difficult to measure these user costs accurately.

²For instance, increases in the U.S. CPI automatically increase federal income tax brackets and some deductions and exemptions as well as trigger boosts in social security benefits, federal employee pensions, and interest payments on inflation-protected securities.

³The issue of whether or not a price inflation objective should be stated as a numerical inflation target is not relevant to our analysis. Our focus is the construction of a measure of underlying inflation that both satisfies some given criteria and is useful for policymakers and private agents concerned with the ongoing path of price changes.

⁴As in Mankikar and Paisley (2002) and Brischetto and Richards (2006), one can alternatively describe the role of a core price measure in terms of distinguishing between relative price movements and changes in underlying inflation.

indicate a lack of concern for the impact of these price changes on current movements in the cost of living.⁶

The development of the core inflation concept appears to have begun in the early 1970s. An early (and likely initial) construction, associated with the late Otto Eckstein, was a weighted growth of unit labor and capital costs for the

The main reason to focus on the behavior of a core price measure is the belief that there is a significant amount of transitory noise in the movement of aggregate consumer prices.

economy as a whole (Eckstein 1981). The more familiar measure of core inflation as aggregate price growth excluding food and energy appears to have been analyzed first in a systematic fashion in a paper by Gordon (1975b). Gordon’s aggregate “‘core’ price equation” was estimated for final sales prices excluding food and energy.⁷

The name “core inflation” then began to be attached to the growth of price measures excluding food and energy. In 1978, the Bureau of Labor Statistics began to report monthly growth of both the CPI and the producer price index excluding food and energy. The Bureau of Economic Analysis also releases data on the monthly growth of the PCE index excluding food and energy as well as the “market-based” PCE index excluding food and energy. An important point concerning the development of these “core inflation” measures is that little or no specific consideration was given for their future use in the formulation of monetary policy.

⁵Some researchers (Aoki 2001; Benigno 2004; Goodfriend and King 1997) have argued that the appropriate goal for monetary policy should be set in terms of a measure of “core” inflation. However, these authors are referring to a measure that comprises “sticky” prices—those prices that change at fixed intervals—and excludes “flexible” prices that may change at any time. The authors’ use of this terminology may stem from the view that their goal inflation series is somewhat comparable to the aggregate index less food and energy, since food and energy prices may be much more flexible than others. Nevertheless, it is important to note that the models underlying this argument are highly stylized. Moreover, while the argument that policy should be concerned primarily with changes in “sticky” prices to offset the resulting inefficiencies may have merits, there is the more overriding concern that it is very difficult to develop an index of “stickiness” to evaluate candidate core inflation measures.

⁶That is, this treatment may alleviate the apparent disconnect and occasional sharp deviation between changes in a cost-of-living index and in the inflation measure(s) that may be the focus of central bank discussion.

⁷In a slightly earlier piece, Gordon (1975a) refers to 1973-74 inflation as comprising several components, including “underlying ‘hard-core’ inflation” (p. 184).

3. CORE INFLATION: PROPOSED MEASURES AND EVALUATION

3.1 Candidate Core Inflation Measures

Although the term “core inflation” has long meant an inflation series excluding food and energy price changes, alternative measures of core inflation have been proposed. This development likely reflects the lack of a widely accepted definition of core inflation. These alternative measures of core inflation are derived using one of two approaches. Borrowing the terminology of Mankikar and Paisley (2002), we refer to these methods as the “statistical approach” and the “model-based approach.”

The statistical approach derives measures of core inflation by performing a predetermined operation on an aggregate price index. The operation may involve excluding certain items from the price index, re-weighting the components of the price index, or smoothing time-series movements in the price index. Alternatively, the model-based approach typically derives measures of core inflation by imposing restrictions from economic theory within the context of a multivariate econometric analysis. This approach leads to estimates of core inflation that may be associated with dynamic factor models or defined as a component of measured inflation possessing particular interactive effects with other variables.⁸

For this study, we restrict our attention to measures of core inflation associated with the statistical approach. We do this for several reasons. One is that the statistical approach yields core inflation measures that are more widely used by central banks and are more familiar to the public. Measures of this type often appear in central bank discussions of monetary policy or in the media. Another reason is that there is little consensus about the specification and identification schemes of model-based measures of core inflation. Last, there is a marked difference between the two approaches in terms of complexity. Model-based core inflation measures could remain problematic to policymakers and the public because the concepts underlying their design can be abstract and their construction computationally demanding. On the contrary, while there is a variety of core inflation measures associated with the statistical approach, each measure is relatively easy to understand and compute.

Within the statistical approach, the core inflation measures reflect very different characterizations of transitory price

⁸As examples, Velde (2006) defines core inflation as the (unobserved) component common to a large number of individual price series, while Quah and Vahey (1995) define core inflation as the component of measured inflation that is uncorrelated with output at medium- to long-run horizons.

movements. Some of these measures associate the bulk of transitory price fluctuations with specific components, thereby prompting their exclusion from an aggregate price index. We consider two examples of this type of core inflation measure. One is based on the conventional practice of excluding food and energy price changes from movements in an aggregate series. The other has been proposed by Clark (2001), who argues for a core measure of inflation that removes only energy price changes from movements in an aggregate series. His motivation is that food prices, at least at the consumer level, likely react to many of the same forces that influence other retail prices, whereas energy price changes are dominated by transitory commodity price shifts.

As an alternative to core inflation measures that remove some prespecified item(s) in every period, Bryan and Cecchetti (1994) advance a measure that involves re-weighting all the

Although the term “core inflation” has long meant an inflation series excluding food and energy price changes, alternative measures of core inflation have been proposed. This development likely reflects the lack of a widely accepted definition of core inflation.

components in the price index. Specifically, their proposed core inflation measure is the weighted median price change in a period, which is defined as the price change in the period for that product such that half the expenditure is for items whose prices are rising just as, or more, rapidly, and half is for items whose price changes are rising just as, or more, slowly. The weighted median is related to the “trimmed mean” concept of core inflation (Dolmas 2005).⁹ Bryan and Cecchetti’s argument for focusing on measures constructed along these lines is that

⁹The trimmed mean is the average price change computed when omitting a specified percentage of the highest and lowest price changes of products (weighted by their expenditure share) for a period. While our analysis does not include the trimmed mean measure among the candidate series, we believe that some caution needs to be exercised in evaluating this measure. Specifically, researchers typically use full-sample estimation techniques to determine how much the distribution of price changes should be trimmed. However, the use of a criterion function to optimally select the amount of trimming could favorably bias the performance of this measure within a particular period of interest. In our view, any evaluation of a trimmed mean measure should be undertaken using recursive estimation so that the trimmed mean is constructed sequentially. This method would circumvent any difficulties that arise from allowing the future history of the data to impact the construct of the series during an earlier time period.

the tails of the price distribution are mainly associated with temporary price level effects; thus, systematically eliminating their influence should yield a more robust measure of the persistent component of inflation.

In contrast to the weighted median that smoothes the cross-section of price changes, Cogley (2002) develops a core measure of inflation that down-weights past changes in the price index. His proposed core inflation measure involves the exponential smoothing of current and past aggregate price changes. The motivation for this measure is the idea that the government and private sector use adaptive methods to learn about a world in which there are occasional regime shifts in mean inflation.

For the analysis, we examine the following four candidate core inflation measures noted above:¹⁰

- the aggregate inflation series excluding food and energy,
- the aggregate inflation series excluding energy proposed by Clark (2001),
- the weighted median measure of the aggregate inflation series proposed by Bryan and Cecchetti (1994),¹¹ and
- the exponentially smoothed version of the aggregate inflation series proposed by Cogley (2002).

Cogley's formulation is given as:

$$(1) \quad \tilde{\pi}_t = g_0 \sum_{j=0}^{\infty} (1 - g_0)^j \pi_{t-j}, \quad 0 < g_0 < 1,$$

where π denotes the relevant aggregate inflation measure. Equation 1 defines the core measure as a one-sided geometric distributed lag of current and past inflation. We follow Cogley and set the gain parameter $g_0 = 0.125$.

3.2 Performance Criteria

Previously, we argued that core inflation should be viewed as an intermediate target for an aggregate inflation goal. Using this proposition as a guide, we evaluate the candidate core measures of inflation based on criteria comparable to those discussed in Wynne (1999):¹²

1. *Transparency of construction.* It may be helpful to build a core price measure in a straightforward, relatively easy

¹⁰In addition to these four series, Rich and Steindel (2005) examine exponentially smoothed versions of the ex food and energy, ex energy, and weighted median series as candidate core inflation measures.

¹¹The Federal Reserve Bank of Cleveland issues monthly estimates of the change in the weighted median CPI. The Bank has recently announced changes in the procedures used to construct this measure (Bryan and Meyer 2007) based on the work of Brischetto and Richards (2006). Our computations are based on the older procedure.

fashion. This criterion facilitates the communication of the concept in the policy dialogue.

2. *Similarity of means.* A core measure should have a mean comparable to the goal inflation series over a long period of time.
3. *Tracking the trend rate of inflation.* A core measure should display a close coherence to the underlying trend in the goal inflation series.
4. *Explanatory content.* A core measure should explain past movements in the goal inflation series as well as provide information about potential future developments.

It is important to note, however, that in the literature there has been little uniformity in the criteria used to evaluate core measures of inflation. For example, Cogley (2002) focuses on the within-sample regression fit of core inflation measures (part of criterion 4). Bryan and Cecchetti (1994) examine the marginal within-sample predictive content of core inflation

It is important to note . . . that in the literature there has been little uniformity in the criteria used to evaluate core measures of inflation.

measures as well as their out-of-sample forecast performance (criterion 4). Clark (2001) judges core inflation measures based on their complexity, similarity of means, ability to track a measure of the trend rate of inflation, and within-sample predictive content (criteria 1, 2, 3, and part of criterion 4). Consequently, our set of criteria listed above is not only consistent with the attributes considered in other studies, but also broader in scope.

Given the lack of a common set of performance criteria for core inflation measures in the literature, a similar issue arises concerning the choice of the goal inflation series. Bryan and Cecchetti (1994), Clark (2001), Cogley (2002), and Blinder and Reis (2005) examine the (standard published) CPI, whereas Dolmas (2005) and Smith (2006) examine the PCE deflator, and Smith (2004) and Khettry and Mester (2006) examine both

¹²Silver (2006) also discusses a wide range of comparable criteria for judging the relative merits of proposed core inflation measures. Wynne (1999), like Bryan and Cecchetti (1994), notes that at times the rationale for the construction of a core price index has been to identify the common component of price changes attributable to monetary policy. If such is the purpose of a core price index, however, then it is not altogether clear why one would confine the measure to elements of household price indexes. The difficulty is that monetary policy affects the demand for all types of products in complex ways. These demand effects are not necessarily similar for household and nonhousehold prices, nor is there any strong reason to assume that the distribution of monetary policy effects between household and other prices will be stable over time.

the CPI and the PCE deflator. One can also extend the list of differences among studies to include sample period, model specification, forecasting horizon, and data frequency. Thus, it seems unlikely under the circumstances that any type of consensus about core inflation measures would emerge, and the divergence of conclusions in the literature bears this out.

In light of the previous discussion, our study attempts to correct for the lack of standardization in the evaluation of core inflation measures. When it is feasible to encompass the features of other studies, such as in the selection of performance criterion and choice of goal inflation series, we do so. When it is not feasible (most notably, across sample period, model specification, forecast horizon, and data frequency), we rely on our judgment to ensure that these dimensions of the evaluation process are reasonable and similar to those adopted in other studies. Although we recognize that there are limits to the generality of our framework, we nevertheless believe that it offers an improved basis for judging the capabilities of core inflation measures, and that it may help to clarify the observed differences across previous studies.

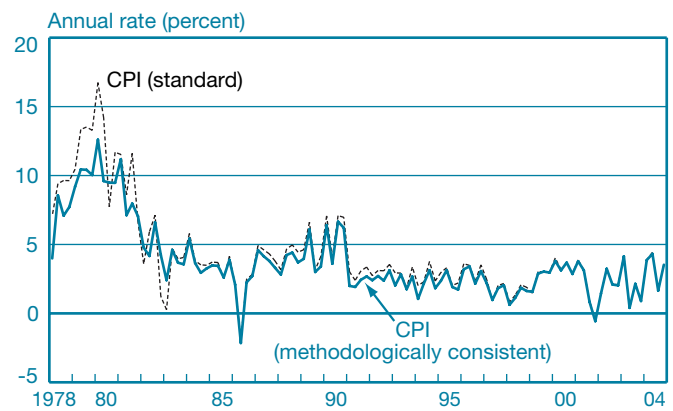
There is one additional point that merits attention, given the conflicting evidence reported in previous studies. Our discussion has emphasized how the design of core inflation measures is influenced by views about the nature of transitory price movements. It is important to note that these views reflect not only an explicit statement about the sources of transitory price movements, but also an implicit assumption concerning the invariance of these sources. If the pattern of transitory price movements were to change over time, then the reliability of core inflation measures would likely be affected. Moreover, if changes in the pattern of transitory price movements were to coincide with different sample periods used for estimation or forecasting purposes, then one would expect to observe variation in the relative performance of the core inflation measures across the sample periods. Keeping these considerations in mind, we now turn to the empirical framework.

4. EMPIRICAL FRAMEWORK

For the analysis, we restrict our attention to aggregate inflation measures that would likely be of interest to policymakers and the public. We select two measures: quarterly growth in the PCE index and quarterly growth in the methodologically consistent CPI. Because the PCE index has gained considerable prominence in U.S. monetary policymaking at the expense of the CPI in recent years, we find that it is instructive to analyze the PCE in parallel with the CPI.¹³

The PCE index is produced while constructing the National Income and Product Accounts data, with the quarterly index

CHART 1
CPI Inflation: 1978-2004



Source: U.S. Department of Labor, Bureau of Labor Statistics.

excluding food and energy starting in 1959.¹⁴ The methodologically consistent CPI is a less familiar price index. It is basically a reconstruction of the CPI designed to, as closely as possible, utilize current procedures to compute the prices of individual products. The major, but by no means sole, difference from the standard CPI is the extension of the current “rental equivalence” method of computing homeowners’ costs to data prior to 1984. The key advantage of using such a series is that it controls for any impact on the statistical results that may arise from changes in the methods used to construct the CPI. A major disadvantage is that the series starts in 1978, limiting the time period available for the analysis.¹⁵

Not surprisingly, the aggregate methodologically consistent CPI series has differed from the standard published CPI series (Chart 1). The divergence is most notable, and quite striking, prior to 1984, when the measurement of homeownership costs in the standard CPI was changed to the owners’ equivalent rent concept used for the entire history of the methodologically consistent series.

We now provide additional details on variable construction, metrics, model specifications, and testing procedures used in our analysis. The criteria used to evaluate the core inflation measures differ in terms of complexity. Whereas comparing the average rates of aggregate inflation and core inflation is relatively straightforward, the same can not be said for tracking the trend rate of inflation and either explaining or forecasting movements in aggregate inflation.

¹³For example, Federal Open Market Committee participants now report their projections of inflation using the PCE index and the PCE index excluding food and energy, rather than the CPI excluding food and energy.

¹⁴We use the vintage of data available before the 2005 annual midyear revision.

¹⁵The methodologically consistent CPI is available monthly from the Bureau of Labor Statistics under the name “CPI-URX,” and is often referred to as the “research series.” We use the more cumbersome title to emphasize its advantages in statistical analysis. Stewart and Reed (1999) describe the index.

4.1 Tracking Trend Inflation

There are two ways to evaluate how well a core inflation measure tracks trend inflation. One way is based on the idea that a core inflation measure should neither understate nor overstate the long-run rate of growth of the goal price index. We can assess this feature by comparing the long-run means of a core inflation measure and aggregate inflation. Another way is based on the idea that a core inflation measure should match the movements in the trend rate of inflation over time. This assessment, however, requires additional assumptions about the way to estimate trend inflation and the metric to gauge the deviation between the series.

To construct the measure of trend inflation, we apply the Baxter-King (1999) band-pass filter to the data.¹⁶ The band-pass filter returns a component that eliminates all periods less than eight years (thirty-two quarters).¹⁷ To gauge the accuracy with which core inflation tracks trend inflation, we follow Clark (2001) and use a measure of volatility for this assessment. Specifically, we compute the root mean squared error (RMSE) of the difference between trend inflation and core inflation:

$$(2) \quad RMSE(\pi^{TREND} - \pi^{CORE}) = \sqrt{\frac{1}{T} \sum_{t=1}^T (\pi_t^{TREND} - \pi_t^{CORE})^2},$$

where π_t^{TREND} is an estimate of the trend of inflation at time t and π_t^{CORE} is a particular measure of core inflation at time t .

4.2 Model Specification and Testing Procedures

Our final set of criteria addresses the ability of the candidate core inflation measures to account for movements in aggregate inflation both within sample and out of sample. The following specification serves as the benchmark model for this part of the analysis:

$$(3) \quad \pi_{t+h} - \pi_t = \alpha_h + \beta_h(\pi_t - \pi_t^{CORE}) + \varepsilon_{t+h},$$

where $\pi_{t+h} = (400) * \ln(P_{t+h}/P_{t+h-1})$ is quarterly inflation h -quarters-ahead reported at an annual rate, $\pi_t = (400) * \ln(P_t/P_{t-1})$ is current quarterly inflation

¹⁶Dolmas (2005) uses the Baxter-King band-pass filter to construct an estimate of trend inflation, while Bryan and Cecchetti (1994) and Clark (2001) use a centered moving average. Recently, Comin and Gertler (2006) have advocated the use of the Baxter-King filter against alternatives to examine movements in aggregate activity.

¹⁷While the band-pass filter is attractive for isolating components with particular periodicities, the estimate of a component at a point in time is based upon both past and future values of a series. Consequently, the band-pass filter is not designed to detect changes in trend inflation in real time and therefore would be of little value to a policymaker.

reported at an annual rate, π_t^{CORE} denotes one of the indicators of current quarterly core inflation measured at an annual rate, and ε_{t+h} is a mean-zero random disturbance term.

The regression model in equation 3 has been used in studies such as Clark (2001), Hogan, Johnson, and Lafèche (2001), Cutler (2001), and Cogley (2002).¹⁸ One attractive feature of the model is its easy interpretation. In particular, the model relates the change in inflation over the next h quarters to the contemporaneous gap between actual inflation and core inflation. That is, the current “core deviation” (transitory movement in inflation) is used to predict how much aggregate inflation will change over the next h quarters. The specification of the model accords with the intuition that if a core measure is identifying current price changes that are expected to dissipate, then the core deviation by definition should be providing a measure of an anticipated reversal in inflation.

Although the formulation of the model in equation 3 is

Our final set of criteria addresses the ability of the candidate core inflation measures to account for movements in aggregate inflation both within sample and out of sample.

admittedly simple, Clark (2001) and others argue that it is consistent with the beliefs of some policymakers and commentators who take movements in core inflation, by themselves, as signals of future changes in inflation. Moreover, the specification of the dependent and independent variables in terms of differences in inflation rates effectively ensures that the two variables are stationary, thereby circumventing any complications arising from the presence of unit roots.¹⁹

Another attractive feature of the model is that we can draw upon the construct of a successful measure of core inflation to obtain restrictions on the parameters in equation 3. In particular, if one adopts the Bryan-Cecchetti (1994) definition of core inflation as “the component of price changes that is expected to persist over medium-run horizons of several years,” then this relationship would imply:

$$(4) \quad \pi_t^{CORE} = E[\pi_{t+h}|I_t],$$

¹⁸Smith (2004, 2006) estimates models that are broadly comparable to equation 3, but her specifications include lagged values of actual inflation and core inflation as additional regressors.

¹⁹During the sample periods considered in this study, U.S. price inflation displays a very high degree of persistence. In particular, it is standard in the literature to model the series as a unit root process.

where E denotes the expectations operator and I_t is an information set that includes information on price changes through time period t .²⁰ From equation 3, the Bryan-Cecchetti definition will hold under the joint restriction $\alpha_h = 0$ and $\beta_h = -1$. The value of β_h is of particular interest because it indicates whether the core deviation is correctly measuring the magnitude of the transitory movement in inflation. Specifically, a value of β_h greater than (less than) unity in absolute value indicates that the measured core deviation understates (overstates) the subsequent changes in inflation, and thus understates (overstates) the magnitude of current transients.

For the within-sample analysis, we undertake the estimation of equation 3, using all available observations over a sample period and allowing the values of h to range from one to twelve quarters. Whenever $h > 1$, there will be overlapping observations caused by the forecast horizon exceeding the sampling interval of the data. Consequently, we use the Newey-West (1987) covariance matrix estimator to account for autocorrelation (and possible conditional heteroskedasticity) of the regression residuals.

In the case of the out-of-sample analysis, we generate the forecasts through recursive estimation of equation 3. Specifically, we use data through quarter $t+h$ to estimate the model, relating the change in the inflation rate between quarter $t+h$ and quarter t to the core deviation in quarter t . Although data on the core deviation for quarter $t+1$ through quarter $t+h$ are not used for estimation, these observations are part of the current information set. Consequently, the estimated model can be iterated forward by h quarters to produce an h -quarters-ahead forecast of inflation $\hat{\pi}_{(t+h)+h}$. We then move the sample forward by one quarter and repeat the exercise. For each measure of core inflation and horizon, the pseudo out-of-sample forecasting procedure will generate a single series of forecast errors $(\pi - \hat{\pi})$.²¹ We can then compute the RMSE of forecasts to compare the performance of the measures of core inflation.²²

The discussion up to this point has considered the core measures in isolation. However, we can augment the regression model to include other variables that may contain additional predictive content for future movements in inflation. In particular, we can extend equation 3 such that:

$$(5) \quad \pi_{t+h} - \pi_t = \alpha_h + \beta_h(\pi_t - \pi_t^{CORE}) + \gamma_h X_t + \varepsilon_{t+h},$$

²⁰For the moment, we can think of I_t as a subset of a larger information set that includes all data available through time t . The information set I_t is merely intended to represent the data used to construct the core measures of inflation examined in this study. As such, it would also include expenditure weights on the various components needed to compute the weighted median measures.

²¹We use the term “pseudo” to acknowledge the fact that the analysis does not use real-time data sets.

²²In another study examining core measures of the methodologically consistent CPI, Marques, Neves, and Sarmento (2003) do not use RMSE to evaluate forecast performance. Instead, they argue that the deviation between aggregate inflation and a measure of core inflation should be correlated with future movements in aggregate inflation but uncorrelated with future movements in the core measure itself.

where X_t is a macroeconomic variable that is taken as an indicator of economic slack.²³ Despite a large set of possible candidates, we restrict ourselves to a measure of the unemployment rate, capacity utilization, and the output gap.²⁴ To adhere to the principle of parsimony, we experiment only with the macroeconomic variables on an individual basis so that γ_h is restricted to be a scalar. However, we consider both the level and first difference of the macroeconomic variables to account for the possibility of rate-of-change effects.

Regarding our evaluation procedure, there are two points that merit special discussion. The first is that we do not propose equation 3 as a model of how actual inflation forecasts are, or should be, prepared. Rather, equation 3 affords us an additional metric to rank the core measures based on their relative accuracy to forecast inflation within the benchmark regression model. A core measure that is better than others on this metric may be considered useful to communicate inflation risks in a straightforward manner, even if a central bank uses other means to construct its internal inflation forecast.

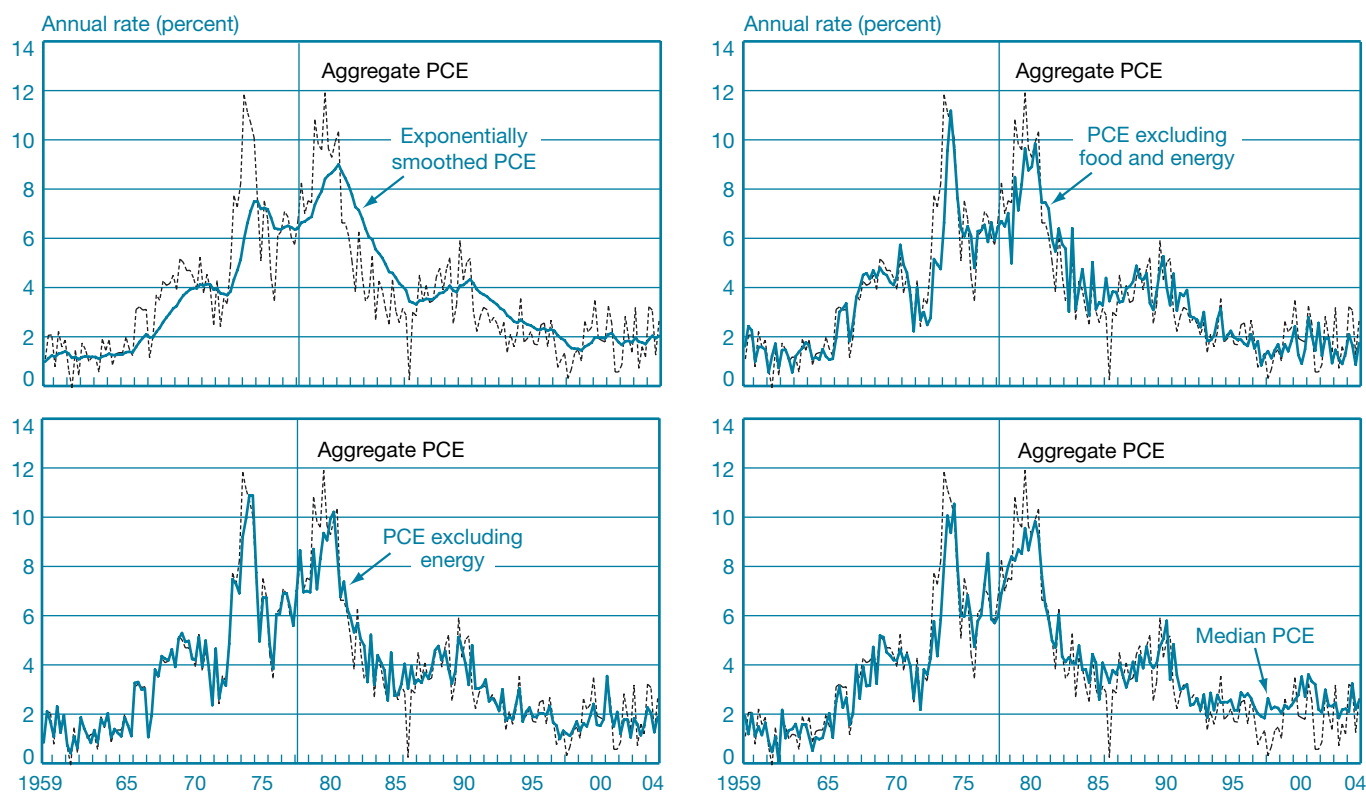
As a second point, we recognize that there may be caveats associated with some of the statistical criteria used to evaluate the candidate core inflation series. For example, the explanatory content of a core measure may vary over time. As documented by Cecchetti, Chu, and Steindel (2000), there is evidence that relationships that appear to satisfactorily predict inflation in one year can often deteriorate in the next. This consideration gains further relevance when the actions of a central bank are taken into account. Specifically, if a core inflation measure was informative about the inflation outlook, then a monetary authority might incorporate this result into its policy formulation. If this changed policy response led to greater stability of inflation, then conventional correlation measures would show a weakening in the link between the goal inflation series and the core inflation measure. Thus, an observed deterioration in the relationship could, paradoxically, result from a core inflation measure remaining a useful indicator of (potential) future developments in inflation.

Concerning this latter point, we examine the issue of stability of the core inflation measures across two dimensions. One dimension focuses on the forecast performance of the core inflation measures over alternative subperiods. The other

²³We could have augmented X_t to include other types of variables such as financial indicators, oil prices, and various types of monetary aggregates. However, we selected measures of (excess) demand pressure in the economy based on the previous results of Cogley (2002) to keep the analysis manageable.

²⁴The unemployment rate is for prime-age males (ages twenty-five to fifty-four) to control for demographic changes. Following Cogley (2002), we construct the output gap measure as $100 * (y_t - \tilde{y}_t)$, where y_t is the log of real GDP and \tilde{y}_t is an estimate of the trend from applying the exponential smoother in equation 1. The measure of capacity utilization is for the manufacturing sector.

CHART 2
PCE Deflator Inflation and Core Measures: 1959-2004



Sources: U.S. Department of Commerce, Bureau of Economic Analysis; authors' calculations.

Note: The vertical bar at 1978:1 indicates the overlap of the various PCE and CPI series during the truncated sample period.

applies structural break tests to investigate the stability of the regression coefficients in equation 3.

5. EMPIRICAL RESULTS

The data on the candidate core measures for PCE inflation start in 1959:2. For the methodologically consistent CPI, the observations on aggregate inflation and the candidate core measures start in 1978:1.²⁵ To provide a basis of comparison and to check the robustness of the results, we also undertake analysis of PCE inflation starting in 1978:1. Because the observations on the weighted median CPI series and the CPI

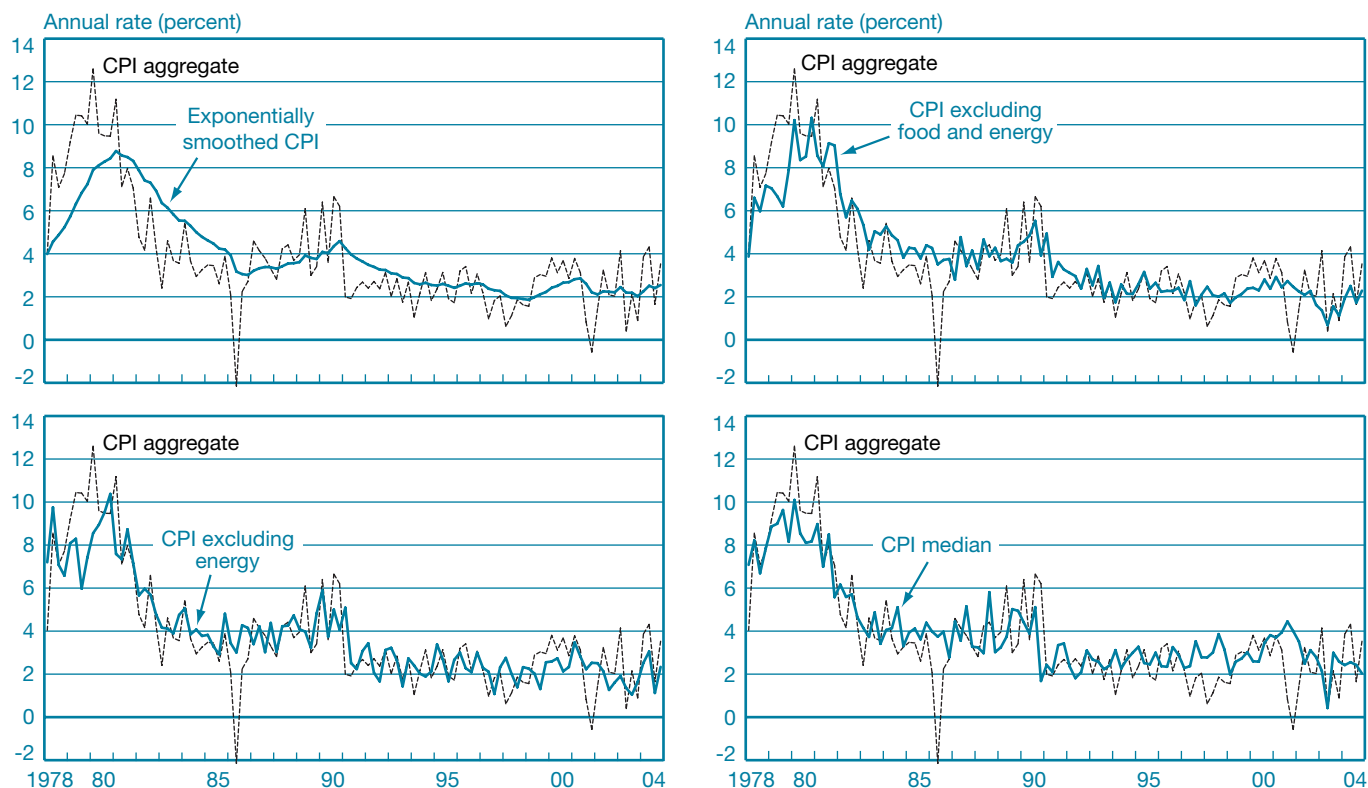
²⁵Quarterly values of the price indexes are averages of the relevant monthly figures. Quarterly growth rates are constructed from these averages. The exception is the growth rate of the CPI ex energy series for 1978:1, which is calculated as the growth of the index from December 1977 to March 1978.

ex energy series end in 2004:4, all of the analyses will end there to maintain consistency.

Chart 2 plots the growth of the PCE index against the candidate core measures over the period 1959-2004, while Chart 3 plots the methodologically consistent CPI over the period 1978-2004. Chart 2 contains a vertical bar at 1978:1 to illustrate the overlap of the various PCE series during the truncated sample period. Although the truncated sample period does not affect the values of actual PCE inflation or the measures excluding energy, excluding food and energy, and the weighted median, it slightly alters the initial values for exponentially smoothed series.²⁶

²⁶Our intention is to treat the PCE index and the CPI index in an identical manner over the truncated sample period. While the change in sample period does not affect the measures of economic slack such as the unemployment rate and capacity utilization, this is not true for our measure of the output gap. As indicated in footnote 24, the exponential smoother is also used to derive the output gap.

CHART 3
CPI Inflation and Core Measures: 1978-2004



Sources: U.S. Department of Labor, Bureau of Labor Statistics; authors' calculations.

5.1 Transparency of Design, Similarity of Means, and Coherence to Trend Inflation

All of our candidate core measures appear to be “transparent” in that they are related to the corresponding measures of aggregate inflation in a relatively straightforward, reproducible fashion.²⁷ Table 1 addresses the next two criteria for a core measure—the ability to match the mean of aggregate inflation and to track the trend rate of inflation. The table provides information on the PCE index for the post-1959 period and on the PCE index and the methodologically consistent CPI for the post-1978 period.

²⁷ Admittedly, the methodologically consistent CPI has not received a great deal of prominence, and the construction of the weighted median and exponentially smoothed measures involves a bit of effort. Because the current expenditure weights of PCE components may be read right off of the corresponding current-dollar consumption data, constructing the weighted median PCE series is a straightforward process.

The first criterion is similarity of means. As shown in Table 1, all the candidate core inflation measures have means quite near to that of aggregate PCE inflation during the full sample period. The situation is somewhat different during the post-1978 period. For both the weighted median series and the exponentially smoothed PCE, the means of the core inflation measures are somewhat higher than those of the respective aggregate inflation series.

Formal statistical tests revealed that the observed differences between long-run growth rates in aggregate prices and the core measures are statistically significant for the weighted median PCE and exponentially smoothed PCE during the 1978-2004 period. It is possible that the general process of disinflation that characterized much of the last quarter-century partly accounts for this finding. For example, the exponentially smoothed series are weighted sums of current and past inflation rates. Consequently, if inflation is generally trending downward, then there should be a tendency for the exponentially

TABLE 1

Average Inflation Rates and Volatilities around Trend Percent

Inflation Measure	1959-2004	1961-2002
	Mean ^a	RMSE($\pi^{TREND} - \pi^{CORE}$) ^b
PCE	3.69	—
PCE ex food and energy	3.60	1.03*
PCE ex energy	3.65	1.02*
Median PCE	3.77	1.00*
ES PCE	3.65	0.78
Inflation Measure	1978-2004	1980-2002
	Mean ^a	RMSE($\pi^{TREND} - \pi^{CORE}$) ^b
PCE	3.61	—
PCE ex food and energy	3.58	0.69
PCE ex energy	3.57	0.67
Median PCE	3.89**	0.79*
ES PCE	3.91**	0.73
Inflation Measure	1978-2004	1980-2002
	Mean ^a	RMSE($\pi^{TREND} - \pi^{CORE}$) ^b
CPI	3.80	—
CPI ex food and energy	3.86	0.73
CPI ex energy	3.80	0.76
Median CPI	4.03	0.89*
ES CPI	3.90	0.68

Source: Authors' calculations, based on data from the U.S. Department of Commerce, Bureau of Economic Analysis, and the U.S. Department of Labor, Bureau of Labor Statistics.

Notes: Test statistics were constructed using the Newey-West (1987) covariance matrix estimator. ES denotes an exponentially smoothed series.

$${}^a H_0: \bar{\pi}^{AGGREGATE} = \bar{\pi}^{CORE}$$

^bThe Diebold-Mariano (1995) test statistic considers the null hypothesis of equal root mean squared error (RMSE) against the alternative hypothesis that the RMSE of a relevant benchmark series is lower.

*Statistically significant at the 5 percent level.

**Statistically significant at the 1 percent level.

smoothed measure of inflation to be a bit higher than current inflation. If inflation is declining, then there may also be a tendency for the outliers in the cross-sectional distribution of price change to be skewed toward negative or low readings, and thus contribute to the weighted median price change running a bit higher than the average price change.

The second criterion is the proximity of a candidate core measure to the current underlying trend in the goal inflation series. Chart 4 depicts the trend estimates, from which we drop two years of data from the beginning and end of the band-pass filtered series because they are relatively poorly estimated. Turning back to Table 1, we note that the RMSEs show that the exponentially smoothed measures tend to track the trend more

closely for CPI inflation and PCE inflation during the longer sample, whereas the ex energy measure tracks the trend in PCE inflation more closely during the shorter sample.

Similar to the analysis comparing long-run growth rates, we can conduct formal tests to determine if the observed differences in RMSE are statistically significant. To address this issue, we construct the Diebold-Mariano (1995) test statistic, which allows us to consider the null hypothesis that two core inflation measures track the trend rate of inflation equally well against the alternative hypothesis that one core inflation measure tracks trend inflation more accurately than the other. For the tests, we select the core inflation measure associated with the lowest RMSE as the benchmark series and then make comparisons with the other core measures on an individual basis. We employ an error criterion for the test based on the difference in squared prediction errors across the two core inflation measures:

$$(6) \quad d_{t+i} = (\hat{\epsilon}_{1,t+i}^2 - \hat{\epsilon}_{2,t+i}^2),$$

where $\hat{\epsilon}_{j,t+i}^2 = (\pi_{t+i}^{TREND} - \pi_{j,t+i}^{CORE})^2$ and d_{t+i} is the differential loss in period $t+i$ using the benchmark series ($j=1$) versus the alternative series ($j=2$). The Diebold-Mariano test essentially determines whether the mean differential loss \bar{d} across a selected sample period is statistically different from 0.

The Diebold-Mariano tests yield similar results during the shorter sample period. The ex energy measure performs significantly better than the weighted median in tracking trend PCE inflation, whereas the exponentially smoothed measure performs significantly better than the weighted median in tracking trend CPI inflation. For the longer sample, however,

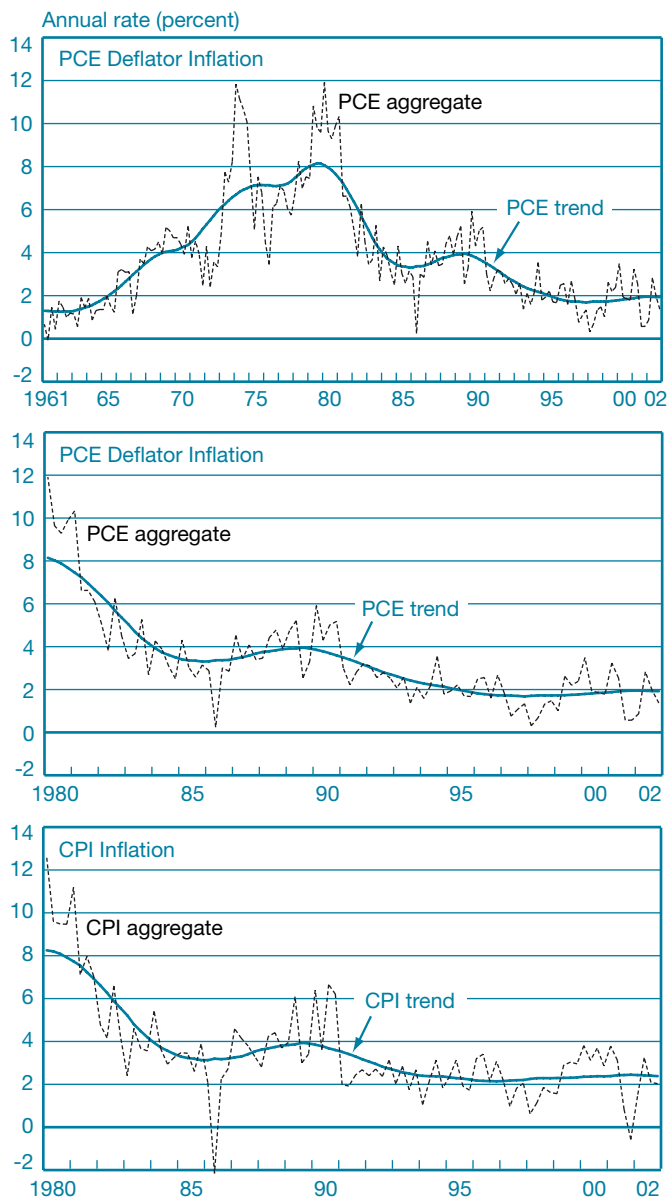
All of our candidate core measures appear to be “transparent” in that they are related to the corresponding measures of aggregate inflation in a relatively straightforward, reproducible fashion.

the exponentially smoothed measure performs significantly better than each alternative core measure in tracking trend PCE inflation. This latter result is driven principally by an episode from the early and mid-1970s in which there was a pronounced deviation between trend PCE inflation and the ex food and energy series, ex energy series, and weighted median series.²⁸

²⁸The reason for the deviation is that trend PCE inflation was much higher than these three core inflation series during this episode.

CHART 4

Aggregate Inflation and Estimated Trend Inflation Using Band-Pass Filter



Sources: U.S. Department of Commerce, Bureau of Economic Analysis; authors' calculations.

An initial read of the findings suggests that the performance of the core inflation measures is fairly similar on balance. While the exponentially smoothed PCE series is much better than all other core inflation measures at tracking trend inflation over the longer sample, its performance does not carry over to either CPI or PCE inflation during the shorter sample. Moreover, the exponentially smoothed PCE series actually overstates the

average rate of PCE inflation during the shorter sample. Even though there is no superior core inflation measure that emerges from Table 1, one could argue that the weighted median series is distinguished by its relatively poor performance in tracking trend inflation. One explanation for this feature of the weighted median may be its re-weighting of the individual price series and thus the weights on their relative trends. Such re-weighting also occurs in the construction of the ex food and energy and the ex energy measures, but it is more extensive for the weighted median.

5.2 Within-Sample Evaluation

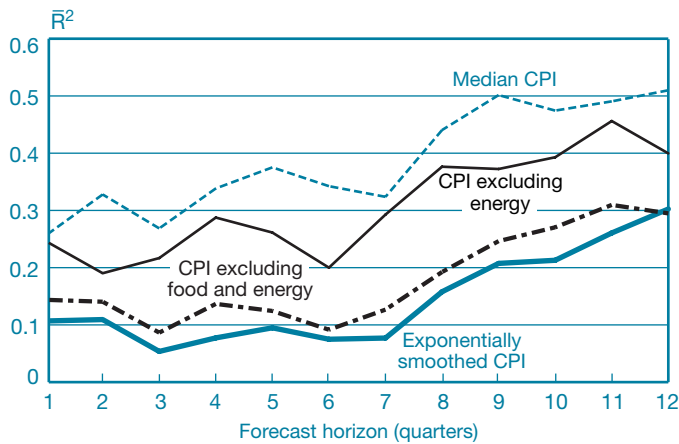
We now examine the results from estimating equation 3 across the various sample periods, forecast horizons, and core measures. Because of the large number of regressions and our desire to conserve space in the reporting of results, we devote the bulk of our discussion to the results using the (methodologically consistent) CPI as the inflation measure, and then summarize the findings for PCE inflation. In a further effort to save space, we generally do not report estimates of individual parameters and standard errors. More complete results may be found in Rich and Steindel (2005).

Chart 5 plots the goodness of fit, as measured by the \bar{R}^2 statistic, for the core inflation measures over horizons that range from one to twelve quarters. Some of the core measures appear to explain a significant amount of CPI growth over these horizons. This is particularly true for the weighted median, which may not fare well in terms of tracking trend inflation but is able to explain approximately 50 percent of the total variation in CPI inflation at horizons exceeding eight quarters (two years). Alternatively, it is interesting to note that the conventional ex food and energy measure and the exponentially smoothed measure tend to have the lowest explanatory content for CPI inflation.

In addition, there is a marked increase in the explanatory content of the core measures as the forecast horizon increases beyond six quarters. This result corroborates the previous findings of Hogan, Johnson, and Lafèche (2001) and Cogley (2002) and is consistent with the intended design of a core measure to identify transients in the data. Not surprisingly, one would expect a greater effectiveness at filtering transitory price changes to translate directly into an improved ability to explain subsequent reversals in inflation.

Table 2 presents tests for unbiasedness ($\alpha_h = 0$ and $\beta_h = -1$). These tests were conducted only at the twelve-quarter horizon to allow sufficient time for the identified transients to dissipate. We strongly reject unbiasedness for the weighted median measure, with the test statistic for the ex energy measure

CHART 5
CPI Inflation: 1978-2004
Within-Sample Fit



Source: Authors' calculations, based on data from the U.S. Department of Labor, Bureau of Labor Statistics.

just slightly below the critical value at the 5 percent significance level.²⁹ It is interesting to note that the source of bias does not stem from the slope coefficient, as all of the core measures satisfy the individual restriction $\beta_{12} = -1$ at conventional significance levels. Rather, the bias is related to the intercept displaying a large and negative value. This result is not particularly surprising, given the behavior of inflation over the post-1980 period. Specifically, it would appear that the core measures are effectively removing the transient noise around trend inflation, although some measures are unable to account for the sustained decline in trend inflation over this sample period.

As discussed previously, there is no reason to confine our attention to measures of core inflation when explaining subsequent movements in aggregate inflation. We now consider the results from estimating equation 5, in which we combine the core measures with macroeconomic variables that are conventionally viewed as indicators of slack in the economy. Because of the even larger number of estimated regressions involved in this part of the analysis, we again elect to provide a summary of the main findings.

These results suggest four main conclusions:

First, the general features associated with the predictability of inflation carry over from the univariate analysis. That is, the \bar{R}^2 for the combinations of core inflation and macroeconomic variables tends to rise as the horizon increases.

Second, although the addition of macroeconomic variables can improve the predictability of aggregate inflation, their contribution can vary considerably across the core measures.

²⁹The test statistic is distributed asymptotically as a chi-square random variable with two degrees of freedom.

TABLE 2
Unbiasedness Test

Core Inflation Measure	α_{12}^a	β_{12}^b	$H_0: \alpha_{12} = 0$ and $\beta_{12} = -1$
CPI ex food and energy	-0.973 (0.507)	-1.024 (0.166)	3.817 <i>p</i> -value = 0.148
CPI ex energy	-0.906* (0.440)	-1.231 (0.153)	5.658 <i>p</i> -value = 0.059
Median CPI	-1.161** (0.411)	-1.332 (0.263)	8.678 <i>p</i> -value = 0.013
ES CPI	-0.936 (0.529)	-0.885 (0.224)	3.139 <i>p</i> -value = 0.208

Source: Authors' calculations, based on data from the U.S. Department of Labor, Bureau of Labor Statistics.

Notes: Standard errors, reported in parentheses, were calculated using the Newey-West (1987) covariance matrix estimator. ES denotes an exponentially smoothed series.

^a $H_0: \alpha_{12} = 0$.

^b $H_0: \beta_{12} = -1$.

*Statistically significant at the 5 percent level.

**Statistically significant at the 1 percent level.

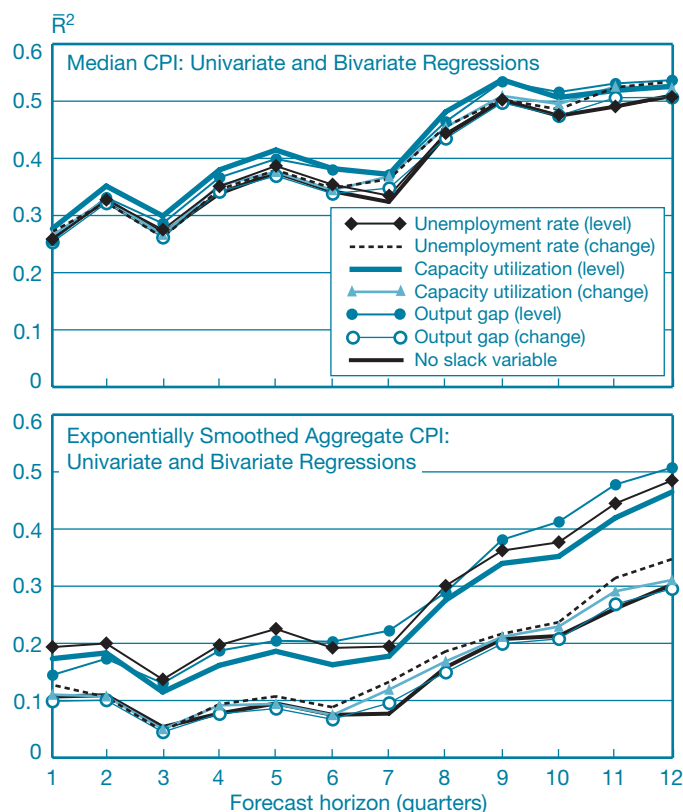
The top and bottom panels of Chart 6 depict the (adjusted) goodness-of-fit measures for both the univariate regressions and bivariate regressions for the weighted median CPI and exponentially smoothed CPI, respectively. As shown in the top panel, the macroeconomic variables offer very little improvement in the fit of the regression over the various horizons.³⁰ In contrast, the bottom panel shows a marked improvement in the fit of the regression when we include capacity utilization, the output gap, or the unemployment rate as a second explanatory variable.

Third, the predictive power of the core measures, in combination with the macroeconomic variables, was almost always higher when the macroeconomic variables entered the regression in level form rather than as a first difference. We interpret this finding as indicating the absence of "rate-of-change" effects. The bottom panel of Chart 6 is representative of these findings. Within the candidate list of macroeconomic variables, the level of the unemployment rate typically resulted in the largest improvement in the fit of the regression equation.

Finally, the highest \bar{R}^2 for the CPI inflation regressions was associated with the ex energy and weighted median measures (along with the level of a macroeconomic variable). This latter finding contrasts sharply with that of Cogley (2002) and likely reflects our use of the methodologically consistent CPI series as well as a difference in sample periods.

³⁰Admittedly, this statement may require some qualification due to the fairly impressive fit of the univariate regression on its own.

CHART 6
CPI Inflation: 1978-2004



Source: Authors' calculations, based on data from the U.S. Department of Labor, Bureau of Labor Statistics.

5.3 Out-of-Sample Forecast Evaluation

For the out-of-sample analysis, we report the results only for horizons corresponding to four quarters, eight quarters, and twelve quarters. These horizons seem to be of particular interest to central banks, and again allow us to keep the discussion manageable.³¹ We select two starting dates for our out-of-sample forecasts: 1990:1 and 1995:1. That is, the first four-quarter out-of-sample forecast corresponds to 1990:4.³² Similar to the within-sample analysis, all out-of-sample forecast analyses end in 2004:4.

³¹Rich and Steindel (2005) also report results for a one-quarter forecast horizon.

³²As we discussed, this forecast is based on model estimation for the four-quarter-ahead change in the inflation rate up through 1989:4 using data on the regressors for 1988:4 and earlier. The forecast is then constructed by iterating the model forward by four quarters and using the available observations on the regressors for 1989:4.

TABLE 3
Forecasting Performance of Alternative Measures of Core Inflation: RMSE of Univariate Forecasts

Core Inflation Measure	Post-1990 Sample			Post-1995 Sample		
	RMSE ($\pi - \hat{\pi}$)					
	<i>h</i> =4	<i>h</i> =8	<i>h</i> =12	<i>h</i> =4	<i>h</i> =8	<i>h</i> =12
CPI ex food and energy	1.523	1.480	1.703**	1.483	1.591	1.885
CPI ex energy	1.432	1.555	1.825**	1.382	1.574	1.867
Median CPI	1.383	1.477	1.868	1.491	1.467	1.658
ES CPI	1.586	1.455	1.506	1.518	1.607	1.752

Source: Authors' calculations, based on data from the U.S. Department of Labor, Bureau of Labor Statistics.

Notes: The Diebold-Mariano (1995) test statistic considers the null hypothesis of equal root mean squared error (RMSE) against the alternative hypothesis that the RMSE of a relevant benchmark series is lower. Test statistics were constructed using the Newey-West (1987) covariance matrix estimator. Figures in bold are the lowest RMSE at each horizon for the various out-of-sample forecast periods. ES denotes an exponentially smoothed series.

*Statistically significant at the 5 percent level.

**Statistically significant at the 1 percent level.

We restrict our discussion to the univariate forecasts based on the regression model in equation 3. Table 3 reports the RMSE statistics, in annual percentage terms, for the post-1990 and post-1995 periods. We highlight in bold the lowest RMSE at each horizon for the various out-of sample forecast periods.

Taken together, the evidence indicates that the forecast performance of the core measures, in both relative and absolute terms, can vary over the choice of sample period. It also varies across forecast horizons, although the differences are more noticeable and emerge more clearly over the medium run of several years. The exponentially smoothed measure delivers the lowest RMSE at the longer horizons in the post-1990 sample. In contrast, in the post-1995 sample, the weighted median measure does well, and the relative performance of the exponentially smoothed measure deteriorates noticeably. For example, the exponentially smoothed measure had the lowest RMSE for the entire post-1990 period at the eight-quarter (two-year) horizon, but the highest RMSE at the same horizon during the post-1995 period. In addition, the ex energy measure performed quite poorly at the longer forecast horizons. This finding also contrasts with Clark's (2001) reported evidence and, again, likely reflects differences in sample periods as well as in our use of the methodologically consistent CPI series.

As before, we can apply Diebold-Mariano tests to assess if the observed differences in RMSE in Table 3 are statistically significant. When we use conventional significance levels, the evidence indicates that there is almost no difference in forecast performance of the core inflation measures across the various horizons during either sample period. The one exception is the exponentially smoothed series during the post-1990 period, which is more accurate than the ex food and energy series and the ex energy series at the twelve-quarter horizon, although this result does not carry over to the weighted median series. This latter finding is somewhat interesting because of the higher RMSE associated with the weighted median, suggesting that it generally compares favorably with the exponentially smoothed series but is occasionally subject to large forecast errors.

Combining the evidence from the within-sample and out-of-sample analyses, we still conclude that there is no clearly dominant core CPI inflation measure. All measures failed to satisfy the test for unbiasedness. Whereas the weighted median possessed superior within-sample explanatory power for future movements in inflation that is also economically significant, it understated the degree of inflation reversal during the 1978-2004 period. In terms of the exponentially smoothed series, a slightly favorable forecast capability was offset by extremely low within-sample explanatory power. With regard to the ex energy series, and especially the ex food and energy series, these measures were not very noteworthy.

5.4 Results for PCE Inflation

The benchmark model for the PCE index was estimated using both long (1959-2004) and short (1978-2004) sample periods. Table 4 summarizes the findings for the within-sample and out-of-sample analyses. We indicate which core inflation measure generates the highest \bar{R}^2 for the benchmark regression model at horizons of four, eight, and twelve quarters, and we also report the range of the within-sample fit across all the core inflation measures. These results are followed by a listing of the core inflation measure that provides the most accurate out-of-sample forecast at these horizons, with the forecasts starting in 1990 and 1995. We also include forecasts starting in 1980 for the PCE deflator over the longer sample period.

Over the long sample period, the core inflation measures showed a remarkably similar pattern in terms of their explanatory content for future inflation. In particular, the within-sample fit of equation 3 was quite low, with the \bar{R}^2 statistic peaking at around 0.20 at the eight-quarter horizon and then remaining fairly constant. With regard to the

TABLE 4

Summary of Within-Sample and Out-of-Sample Analyses: Best-Performing Measures of Core Inflation

Within-Sample Fit and Range of \bar{R}^2	PCE: 1959-2004	PCE: 1978-2004	
$h=4$	Ex energy (0.014 - 0.081)	Median (0.043 - 0.173)	
$h=8$	Ex energy (0.102 - 0.185)	Median (0.142 - 0.345)	
$h=12$	ES (0.142 - 0.177)	Median (0.252 - 0.366)	
Out-of-Sample RMSE ($\pi - \hat{\pi}$)	Horizon	PCE: 1959-2004	PCE: 1978-2004
1980-2004	$h=4$	Ex energy	—
	$h=8$	Ex energy	—
	$h=12$	Ex energy	—
1990-2004	$h=4$	Ex energy	Ex energy
	$h=8$	Ex food and energy	Median
	$h=12$	Ex energy	Median
1995-2004	$h=4$	Ex energy	Median
	$h=8$	ES	Median
	$h=12$	Ex food and energy	Median

Source: Authors' calculations, based on data from the U.S. Department of Commerce, Bureau of Economic Analysis.

Note: RMSE is root mean squared error; ES denotes an exponentially smoothed series.

parameter restrictions $\alpha_h = 0$ and $\beta_h = -1$, we were unable to reject them either individually or jointly at the twelve-quarter horizon. In contrast to the results for the CPI, we found that all of the core PCE inflation measures possess the property of unbiasedness over the long sample period.

Over the short sample period, the within-sample fit of equation 3 tended to increase with the horizon and was broadly comparable, albeit a bit lower, to those for the CPI regressions. Similar to our findings for the CPI, the weighted median had the most explanatory content for overall inflation across the horizons, with the exponentially smoothed series and ex energy series displaying the least explanatory content. The inclusion of (the level of) macroeconomic variables generally improved the within-sample fit of equation 5 for each core measure, although the magnitude of the improvement was more modest compared with that shown over the long sample. In fact, the \bar{R}^2 of the bivariate regression for each measure of core PCE inflation now displays a greater similarity across the two

sample periods. As was the case for the CPI, the short sample period reveals some evidence of bias in the core inflation measures at the twelve-quarter horizon. We strongly reject unbiasedness for the weighted median, with the test statistics for the ex energy and exponentially smoothed measures just slightly below the critical value at the 5 percent significance level. Moreover, the estimated intercept for all four measures was large, negative, and statistically significantly different from zero at conventional significance levels.

Out-of-sample forecast results (including those starting in 1980 for the PCE regression estimated over the post-1959 period) again varied substantially. When the benchmark model was estimated with data starting in 1959, the ex energy measure typically provided the most accurate forecasts over the post-1980 and post-1990 periods. From a statistical standpoint, however, the relative forecast performance of the ex energy measure was most meaningful at the four-quarter horizon and then tended to dissipate as the forecast horizon increased. In the case of the forecasts for the post-1995 period, the core inflation measure associated with the lowest RMSE varied from the ex energy series (four-quarter horizon) to the exponentially smoothed series (eight-quarter horizon) to the ex food and energy series (twelve-quarter horizon). In most cases, the difference in RMSE across the core inflation measures was not statistically significant.

Conversely, we observed a much more consistent pattern when we examined the out-of-sample forecast results based on estimating the benchmark model with data starting in 1978. For the post-1990 period, the ex energy series produced the lowest RMSE at the four-quarter horizon. In all other cases, however, the weighted median measure produced the most accurate forecasts. Moreover, the forecast performance of the weighted median relative to the other core inflation measures was statistically meaningful at both the eight- and twelve-quarter horizons over the post-1995 period. It is also worth noting that, for each forecast horizon over the post-1990 and post-1995 periods, the ex food and energy series produced the highest RMSE in five out of six possible cases.

The findings for the PCE deflator as well as those presented earlier for the CPI offer little, if any, compelling evidence in support of a preferred measure of core inflation. In many instances, the performance of the candidate series looks roughly comparable. On those occasions in which one series displays superior performance, it sometimes does not carry over to changes in the measure of aggregate inflation or the periods used for estimation and forecasting purposes. When there is evidence indicating that a core inflation measure may be well suited for performing a particular task, the same measure often displays inferior performance in terms of other criteria.

5.5 A Closer Look at the Results: Instability in the Core Inflation Measures, or Something Else?

Even though our findings contrast with claims, either explicit or implicit, in support of a particular measure of core inflation, other studies have drawn conclusions similar to ours. Using Canadian price data, Hogan, Johnson, and Lafèche (2001) examine several proposed measures of core inflation. They find that the candidate series are quite similar in many respects, but there is also evidence that some measures fare better than others along different dimensions. Rather than selecting one measure to perform the role of core inflation, they recommend examining a limited number of measures and using the varied information to assess the inflation outlook. Mankikar and Paisley (2002), examining price data for the United Kingdom, also find that there is no single measure of core inflation that performs well across a set of performance criteria.

Given our inability to identify a preferred core inflation measure, a reasonable reaction might be to look for instability in the relationship between aggregate inflation and the core inflation measures. We alluded to this concern previously in the context of a monetary authority that changed its behavior by incorporating information from a core inflation measure to set future policy.³³ In addition, some commentators have claimed that a wide range of inflation forecasting models experienced a breakdown during the mid-1990s. Further, a comparison of the goodness-of-fit and RMSE results reported for PCE inflation across the longer and shorter samples would appear to hint at the possibility of parameter instability.

To explore this issue more formally, we apply the predictive tests developed by Ghysels, Guay, and Hall (1998) to investigate the stability of the benchmark model. This testing procedure is particularly attractive because it does not require the researcher to specify the break point a priori, but rather allows the data to determine the break point. Because the disparity in the performance of the core measures is more notable at longer horizons, we restrict our attention to horizons of $h=8$ and 12 quarters.³⁴

Details of the test results can be found in Rich and Steindel (2005). In summary, the parameters of equation 3 for PCE inflation over the longer sample period appear stable using

³³Of course, the issue of stability of these types of measures would be relevant in any situation where the economy undergoes a significant structural change.

³⁴We are grateful to Arturo Estrella and Tony Rodrigues for providing the computer program used for this testing procedure. The joint tests of parameter stability impose a common break date across the parameters. The testing procedure also requires that we exclude from consideration a fraction of the sample at each end, so that the set of possible break points lies within an interior range. Consequently, our testing procedure is based on a 5 percent trimming of the sample. Because of the overlapping nature of the data, we again employ the Newey-West (1987) covariance matrix estimator in the construction of the test statistics.

conventional significance levels. For the post-1978 period, the test statistics for PCE inflation also provide little evidence of parameter instability. For CPI inflation, the results speak to more concern about parameter instability, but the evidence is still not very compelling in terms of incidence and statistical significance. Overall, it does not appear that instability in the structure of the underlying estimating model contributes significantly to our results.

As an alternative to the idea of a breakdown in the relationship between aggregate inflation and the core measures, we offer a more intrinsic explanation for our inability to identify a preferred core inflation measure. Simply stated, it is our view that there is too much variability in the nature and sources of transitory price movements to be effectively captured through the design of any one of the core inflation measures. For example, although there may normally be large transitory elements in movements in energy prices or

Simply stated, it is our view that there is too much variability in the nature and sources of transitory price movements to be effectively captured through the design of any one of the core inflation measures.

food prices or both, the sustained run-up over the last few years in oil (and corn) prices highlights episodes in which these movements have displayed a greater degree of persistence. Similarly, the validity of core inflation measures based on a truncation of the distribution of price changes relies on the premise that large price changes—either increases or decreases—in individual items are generally associated with temporary price level effects.³⁵ As Mankikar and Paisley (2002) note, however, theoretical arguments based on menu cost models and staggered-price-setting models argue that both the excluded tails and the included portion of the price-change distribution contain a mixture of transitory and permanent shocks. Likewise, it is entirely reasonable to argue that there have been regime shifts in the mean of U.S. inflation that would justify a down-weighting of past price changes along the lines of the exponentially smoothed measure. However, the usefulness of this particular core inflation measure would seem less compelling during sustained periods of time characterized by intra-regime rather than inter-regime movements in inflation.

³⁵The validity of this type of core inflation measure also depends on the ability to specify the threshold defining the magnitude of a large price change.

While a more detailed characterization of transitory price movements is beyond the scope of this article, we suggest that the lack of a consistent pattern to the movements' behavior could also help account for the differential performance of the core inflation measures documented in other studies. If there is variation in the nature and sources of transitory price movements, then different choices of models, sample periods, and criteria would be expected to yield different results. Thus, we would argue that the diversity of findings simply reflects the confluence of the lack of a consistent pattern to temporary price movements and the lack of standardization in the evaluation process.

6. CONCLUSION

Viewing the stabilization of CPI or PCE inflation as plausible goals for U.S. monetary policy, we evaluate several proposed measures of “core” inflation. Other studies have addressed this issue, but there has been little commonality in their underlying approaches. Thus, a key feature of our analysis is to provide a more consistent basis on which to judge the performance of the candidate core inflation series. This consideration guided us in our selection of model specification, criteria, and sample periods.

Given our empirical framework and greater standardization in the evaluation process, one possible outcome from our study was that a single core inflation measure would emerge as dominant in its performance. This, however, was not the case. Rather, we documented considerable variation in the performance of the candidate series. Further, we noted the rather unremarkable performance of the conventional ex food and energy series. Consequently, the general practice of identifying “core inflation” with an ex food and energy series instead of an alternative series does not seem to be justified based on our analysis.

Although the results of this study do not rule out the potential usefulness of core inflation measures, there appear to be difficulties associated with how best to employ the current set of candidates. One possibility is to weight various criteria and then select the core inflation measure that yields the best performance. However, this approach would be influenced by the highly subjective process of ranking the importance of the criteria. Another possibility is to acknowledge that different core inflation measures seem better suited to performing different tasks, and then adopt the appropriate core inflation measure as the guide for a particular stated purpose. However, this approach would introduce the inconvenience of keeping track of a variety of core inflation measures; moreover, in the

policy area, it could require that a central bank provide the public with a clear understanding of each series. Finally, a central bank could consider the adoption of a model-based measure of core inflation. However, this approach would then

face the previously cited difficulties of choosing the particular definition of core inflation and communicating it to the public. Taken together, these considerations and our results present challenging avenues for future research.

REFERENCES

- Aoki, K. 2001. "Optimal Monetary Policy Response to Relative Price Changes." *JOURNAL OF MONETARY ECONOMICS* 48, no. 1 (August): 55-80.
- Baxter, M., and R. G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *REVIEW OF ECONOMICS AND STATISTICS* 81, no. 4 (November): 575-93.
- Benigno, P. 2004. "Optimal Monetary Policy in a Currency Area." *JOURNAL OF INTERNATIONAL ECONOMICS* 63, no. 2 (July): 293-320.
- Blinder, A. S., and R. Reis. 2005. "Understanding the Greenspan Standard." In *THE GREENSPAN ERA: LESSONS FOR THE FUTURE*. Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 25-27.
- Brischetto, A., and A. J. Richards. 2006. "The Performance of Trimmed Mean Measures of Underlying Inflation." Reserve Bank of Australia Research Discussion Paper no. 2006-10, December.
- Bryan, M. F., and S. G. Cecchetti. 1994. "Measuring Core Inflation." In N. G. Mankiw, ed., *MONETARY POLICY: NBER STUDIES IN BUSINESS CYCLES*, vol. 29, 195-215. Chicago and London: University of Chicago Press.
- Bryan, M. F., and B. H. Meyer. 2007. "Methodological Adjustments to the Median and 16 Percent Trimmed-Mean CPI Estimates." Federal Reserve Bank of Cleveland. Available at <<http://www.clevelandfed.org/research/inflation/US-Inflation/revmcp.pdf>>.
- Cecchetti, S. G., R. S. Chu, and C. Steindel. 2000. "The Unreliability of Inflation Indicators." Federal Reserve Bank of New York *CURRENT ISSUES IN ECONOMICS AND FINANCE* 6, no. 4 (April).
- Clark, T. E. 2001. "Comparing Measures of Core Inflation." Federal Reserve Bank of Kansas City *ECONOMIC REVIEW* 86, no. 2 (second quarter): 5-31.
- Cogley, T. 2002. "A Simple Adaptive Measure of Core Inflation." *JOURNAL OF MONEY, CREDIT, AND BANKING* 34, no. 1 (February): 94-113.
- Comin, D., and M. Gertler. 2006. "Medium-Term Business Cycles." *AMERICAN ECONOMIC REVIEW* 96, no. 3 (June): 523-51.
- Cutler, J. 2001. "Core Inflation in the U.K." Bank of England External MPC Unit Discussion Paper no. 3, March.
- Diebold, F. X., and R. Mariano. 1995. "Comparing Predictive Accuracy." *JOURNAL OF BUSINESS AND ECONOMIC STATISTICS* 13, no. 3 (July): 253-63.
- Dolmas, J. 2005. "Trimmed Mean PCE Inflation." Federal Reserve Bank of Dallas Working Paper no. 0506, July.
- Dotsey, M., and T. Stark. 2005. "The Relationship between Capacity Utilization and Inflation." Federal Reserve Bank of Philadelphia *BUSINESS REVIEW*, second quarter: 8-17.
- Eckstein, O. 1981. *CORE INFLATION*. Englewood Cliffs, N.J.: Prentice-Hall.
- Ghysels, E., A. Guay, and A. Hall. 1998. "Predictive Tests for Structural Change with Unknown Breakpoint." *JOURNAL OF ECONOMETRICS* 82, no. 2 (February): 209-33.
- Goodfriend, M. S., and R. G. King. 1997. "The New Neoclassical Synthesis and the Role of Monetary Policy." In B. S. Bernanke and J. Rotemberg, eds., *NBER MACROECONOMICS ANNUAL 1997*, 231-83. Cambridge: MIT Press.
- Gordon, R. J. 1975a. "Alternative Responses of Policy to External Supply Shocks." *BROOKINGS PAPERS ON ECONOMIC ACTIVITY*, 1975, no. 1: 183-206.
- . 1975b. "The Impact of Aggregate Demand on Prices." *BROOKINGS PAPERS ON ECONOMIC ACTIVITY*, 1975, no. 3: 613-70.
- Hogan, S., M. Johnson, and T. Laflèche. 2001. "Core Inflation." Bank of Canada Technical Report no. 89, January.
- Khettry, N. K., and L. J. Mester. 2006. "Core Inflation as a Predictor of Total Inflation." Federal Reserve Bank of Philadelphia *RESEARCH RAP*, Special Report, April 26.
- Mankikar, A., and J. Paisley. 2002. "What Do Measures of Core Inflation Really Tell Us?" Bank of England *QUARTERLY BULLETIN*, winter: 373-83.
- Marques, C. R., P. D. Neves, and L. M. Sarmento. 2003. "Evaluating Core Inflation Indicators." *ECONOMIC MODELLING* 20, no. 4 (July): 765-75.

REFERENCES (CONTINUED)

- Newey, W. K., and K. D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *ECONOMETRICA* 55, no. 3 (May): 703-8.
- Quah, D., and S. V. Vahey. 1995. "Measuring Core Inflation." *ECONOMIC JOURNAL* 105, no. 432 (September): 1130-44.
- Rich, R., and C. Steindel. 2005. "A Review of Core Inflation and an Evaluation of Its Measures." Federal Reserve Bank of New York *STAFF REPORTS*, no. 236, December.
- Silver, M. 2006. "Core Inflation Measures and Statistical Issues in Choosing among Them." IMF Working Paper no. WP/06/97, April.
- Smith, J. K. 2004. "Weighted Median Inflation: Is This Core Inflation?" *JOURNAL OF MONEY, CREDIT, AND BANKING* 36, no. 2 (April): 253-63.
- . 2006. "PCE Inflation and Core Inflation." Unpublished paper, Lafayette College, July 6.
- Stewart, K. J., and S. B. Reed. 1999. "CPI Research Series Using Current Methods, 1978-98." *MONTHLY LABOR REVIEW* 122, no. 6 (June): 29-38.
- Velde, F. 2006. "An Alternative Measure of Core Inflation." Federal Reserve Bank of Chicago *ECONOMIC PERSPECTIVES* 30, first quarter: 55-65.
- Wynne, M. 1999. "Core Inflation: A Review of Some Conceptual Issues." Federal Reserve Bank of Dallas Working Paper no. 9903, June.

The views expressed are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. The Federal Reserve Bank of New York provides no warranty, express or implied, as to the accuracy, timeliness, completeness, merchantability, or fitness for any particular purpose of any information contained in documents produced and provided by the Federal Reserve Bank of New York in any form or manner whatsoever.